Figure 1: cDNA sequence (and encoded amino acid sequence) of Cynomologous monkey Cathepsin S

```
   1  AGTTGAACTGAAATCTCCCTGCTGCTGCTTTGAAATCTTAGAAGAGAGCCCATCAATTCA
  61  AGGATTCTTACTGTAGGAGCACCTGCTGGTTCTATCACAATGAAACAGCTGGTTTGTGTG
                                          M  K  Q  L  V  C  V
                                          ♦
 121  CTCTTGGTGTGCTCCTCTGCAGTGGCGCAGTTGCATAAAGATCCTACCCTGGATCATCAC
       L  L  V  C  S  S  A  V  A  Q  L  H  K  D  P  T  L  D  H  H
 181  TGGCATCTCTGGAAGAAAACCTATGGCAAACAATACAAGGAAAAGAATGAAGAAGCAGTA
       W  H  L  W  K  K  T  Y  G  K  Q  Y  K  E  K  N  E  E  A  V
 241  CGACGTCTCATCTGGGAAAAGAATCTAAAGTTTGTGATGCTTCATAACCTGGAGCATTCA
       R  R  L  I  W  E  K  N  L  K  F  V  M  L  H  N  L  E  H  S
 301  ATGGGAATGCACTCATATGATCTGGGCATGAACCACCTGGGAGACATGACCAGTGAAGAA
       M  G  M  H  S  Y  D  L  G  M  N  H  L  G  D  M  T  S  E  E
 361  GTGATGTCTTTGATGAGTTCCCTGAGAGTTCCCAGCCAGTGGCAGAGAAATATCACATAT
       V  M  S  L  M  S  S  L  R  V  P  S  Q  W  Q  R  N  I  T  Y
                               ↓
 421  AAGTCAAACGCTAATCAGATATTGCCGGATTCTGTGGACTGGAGAGAGAAAGGGTGTGTT
       K  S  N  A  N  Q  I  L  P  D  S  V  D  W  R  E  K  G  C  V
 481  ACTGAAGTGAAATATCAAGGTTCTTGTGGTGCTTGCTGGGCTTTCAGTGCTGTGGGGGCC
       T  E  V  K  Y  Q  G  S  C  G  A  C  W  A  F  S  A  V  G  A
 541  CTGGAAGCACAGCTGAAGCTGAAAACAGGAAAGCTGGTGTCTCTCAGTGCCCAGAACCTG
       L  E  A  Q  L  K  L  K  T  G  K  L  V  S  L  S  A  Q  N  L
 601  GTGGATTGCTCAACTGAAAAATATGGAAACAAAGGCTGCAATGGTGGCTTCATGACAAGG
       V  D  C  S  T  E  K  Y  G  N  K  G  C  N  G  G  F  M  T  R
 661  GCTTTCCAGTACATCATTGATAACAACGGCATCGACTCAGACGCTTCCTATCCCTACAAA
       A  F  Q  Y  I  I  D  N  N  G  I  D  S  D  A  S  Y  P  Y  K
 721  GCCACGGATCAGAAGTGTCAATATGACTCAAAATATCGTGCTGCCACATGTTCAAAGTAT
       A  T  D  Q  K  C  Q  Y  D  S  K  Y  R  A  A  T  C  S  K  Y
 781  ACTGAACTTCCTTATGGCAGAGAAGATGTCCTGAAAGAAGTTGTGGCCAATAAAGGCCCA
       T  E  L  P  Y  G  R  E  D  V  L  K  E  V  V  A  N  K  G  P
 841  GTGTCTGTTGGTGTGGATGCGAGTCATCCTTCTTTCTTCCTCTACAGAAGTGGTGTCTAC
       V  S  V  G  V  D  A  S  H  P  S  F  F  L  Y  R  S  G  V  Y
 901  TATGAACCATCCTGTACTCAGAATGTGAATCATGGTGTACTTGTGGTTGGCTATGGTGTT
       Y  E  P  S  C  T  Q  N  V  N  H  G  V  L  V  V  G  Y  G  V
 961  CTTAACGGGAAAGAATACTGGCTTGTGAAAAACAGCTGGGGCCGCAACTTTGGTGAAGAA
       L  N  G  K  E  Y  W  L  V  K  N  S  W  G  R  N  F  G  E  E
1021  GGATATATTCGGATGGCAAGAAATAAAGGAAATCATTGTGGGATTGCTAGTTTCCCCTCT
       G  Y  I  R  M  A  R  N  K  G  N  H  C  G  I  A  S  F  P  S
1081  TACCCAGAAATCTAGAGAGGATCTCTTCTTTTTATAACAAATCAAGAAAATATGAAGCAC
       Y  P  E  I  *                                          SEQ ID NO:2
1141  TTTCTCTTAACTTAATTTTTCCTGCTGTATCCAGAAGAAATAATTGTGTCACGATTAATG
1201  TGTATTTACTGTACTAATTAAAAAATATAGTTTGAGGCCGGGCACGGTGGCTCACGCCTG
1261  TAATCCCAGTACTTTGGGAGGCCAAGGCAGGCATATCAACTTGAGGCCAGGAGTTAAAGA
1321  CAGGCCTGGCTAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAACATTAGCGGAGC
1381  GTAATGGTGCATGCCTGTAATCCCAGCTACTTGGGAGGCTGAAGCACAAGATTCCTTGAA
1441  CCCAAGAGGTTGAGGCTGTGGTGAGCTGAGACCACACCACTGTACTCCAGCCTGGACAAC
1501  AGAGTGGAGACTCTGTTTCAAAAAAACAGAAAAGATAATATAGTTTGATTCTTCGTTTTT
1561  TAAAATTTGCAAACCTCAGGAAAAAGTTTGCTAAGTAAATTAGTAGTGTACTATAGATAT
1621  AACTGTATAAAAATTGTTCAACCTAAAACAATCTGTCATTGTTCATTGTTTTATTTTATA
1681  CTCTTTGTCTTTTTAAGACCCCTGATAGCCTTTTGTAACTTGATGGCTTAAAAGTACTTA
1741  ATAAATCTGCCATTTCAAATTTCATTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA     SEQ ID NO:1
```

♦ = cleavage site for pre-sequence

↓ = cleavage site for pro-sequence

* = stop codon

FIGURE 1

Figure 2: Alignment of cathepsin S amino acid sequences from Cynomologous monkey (SEQ ID NO:2, as shown in Fig. 1), human (Genbank accession number BC002642.2), monkey sml cells (*Saimiri boliviensis*), dog (*Canis familiaris*; Genbank accession number AY156692), and mouse (*Mus musculus*; Genbank gi3850787), followed by a consensus sequence.

"cyno cats preprotein" = Cynomologous monkey = SEQ ID NO:2
"smlmonkey cats" = monkey sml cells (*Saimiri boliviensis*) = SEQ ID NO:3
"hcats prepropro.txt" = human (*Homo sapiens*) = SEQ ID NO:4
"dog cats" = dog (*Canis familiaris*) = SEQ ID NO:5
"mouse cats prepro" = mouse (*Mus musculus)* = SEQ ID NO:6
(consensus sequence = SEQ ID NO:7)

FIGURE 2A

ClustalW Formatted Alignments

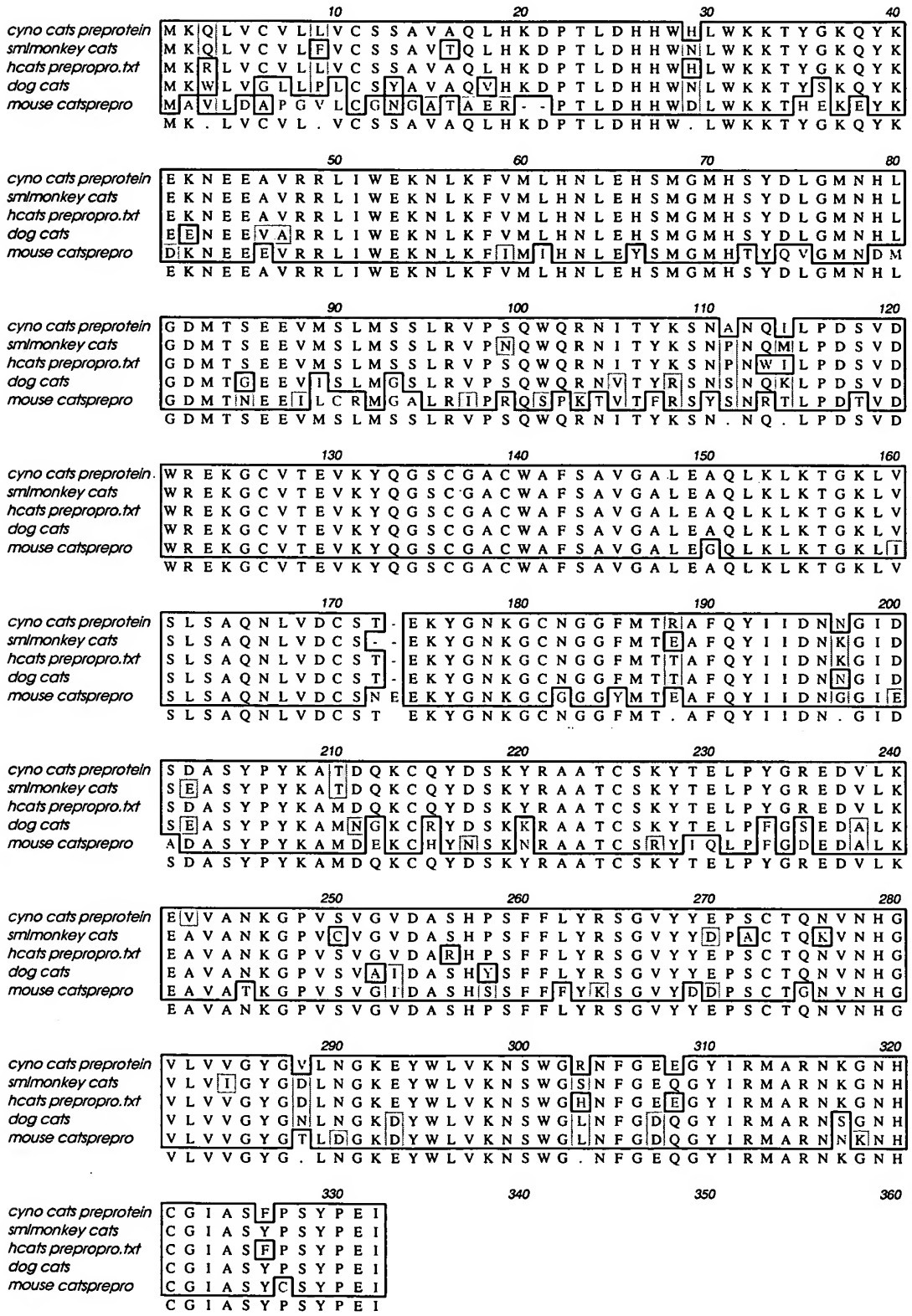

FIGURE 2B

Figure 3: Purification of monkey Cathepsin S (Gelcode stained gel)

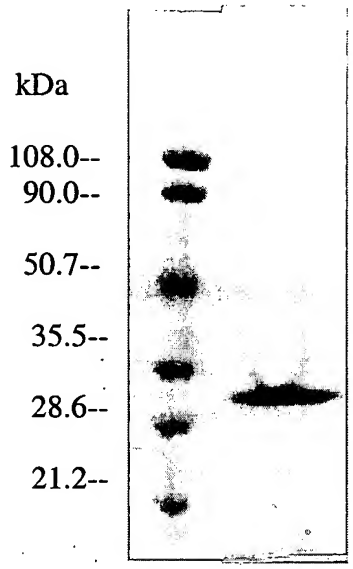

FIGURE 3